UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/801,285	03/07/2001	Jun Matsumoto	7217/64056	3773

7590 08/03/2004

Jay H. Maioli
Cooper & Dunham LLP
1185 Avenue of the Americas
New York, NY 10036

EXAMINER	
LAO, TIM P	

ART UNIT	PAPER NUMBER
2655	9

DATE MAILED: 08/03/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No. 09/801,285	Applicant(s) MATSUMOTO ET AL.	
	Examiner Tim Lao	Art Unit 2655	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *27 May 2004*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1,3-7,9-14,16 and 17* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1,3-7,9-14,16 and 17* is/are rejected.

7)☒ Claim(s) *1, 7, and 14* is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *27 May 2004* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

Response to Arguments

1. Applicant's arguments with respect to claims 1-17 have been considered but are moot in view of the new ground(s) of rejection.

2. The indicated allowability of claims 2-4, 8-11, and 15-16 is withdrawn in view of the newly discovered reference(s) to W. Verhelst et al., "An overlap-add technique based on waveform similarity (WSOLA) for high quality time-scale modification of speech," ICASSP '93. Rejections based on the newly cited reference(s) follow.

Claim Objections

3. Claims 1, 7, and 14 are objected to because of the following informalities:

Regarding claim 1, "(L m)" should be changed to -- (L-m) -- in the "second dividing means ..." limitation.

Regarding claim 7, "(L m)" should be changed to -- (L-m) -- in the "dividing means ..." limitation.

Regarding claim 14, "(L m)" should be changed to -- (L-m) -- in the "dividing a signal ..." limitation.

Appropriate correction is required.

Claim Rejections - 35 USC § 103

4. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

5. Claims 1, 3-7, 9-14, 16, and 17 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kleijn (U.S. Patent 5,517,595, hereinafter "Kleijn") in view of Verhelst et al. ("An overlap-add technique based on waveform similarity (WSOLA) for high quality time-scale modification of speech," IEEE ICASSP '93, hereinafter "Verhelst").

Claim(s)	
1	<u>Kleijn shows:</u> an audio signal processing apparatus (speech coding apparatus: col.2, L.36-37) for reproducing an audio signal by decoding encoded predictive (LP residual, Fig.10: 203; Fig.2; col.4, L.19-20) residual signals produced by forward prediction on a frame by frame basis, the apparatus comprising: excitation source modifying means (Fig.10: 231) for extending or shortening said predictive residual signals on a time axis; and (col.2, L.41-42; col.4, L.57-65) *{1. Excitation source modifying means comprise performing pitch detection and extracting prototype waveform from the linear predictive (LP) residual signal. (col.2, L.41-42)* *2. In extracting prototype waveforms, the residual signal segment is extended for at least one-half pitch period. (col.4, L.57-65)* *3. Forward prediction is e.g., extending the residual signal (periodic for voiced signal) for prediction using past signal samples.}* synthesizing means (Fig.11: 321, 322, 323, 303) for synthesizing (through the LP synthesis filter) an audio signal based on predictive residual signals (the reconstructed residual signal) converted by said excitation source modifying means, (col.2, L.42-L.48) <u>Kleijn does not show:</u> wherein said excitation source modifying means comprises:

first dividing means for dividing said predictive residual signals into a plurality of sub-frames based on a pitch;

second dividing means for dividing a signal of sub-frames into a first signal having a length m, where m is an integer and m<L, where L is the length of said sub-frame, and a second signal having a length (L-m) as a reference signal;

finding means for finding a signal closest to said reference signal from an other sub-frame,

wherein said excitation source modifying means shortens said predictive residual signals by concatenating the first signal and the closest signal.

Verhelst teaches:

first dividing means for dividing said a signal (e.g., x(n)) into a plurality of sub-frames (e.g., segments (1), (1'), (2), (2')) based on a pitch (e.g., $L_k - L_{k-1}$); (see Fig.2 on p.II-556)
{ $L_k - L_{k-1}$ is the local pitch period. (p.II-555, col.2, 1st ¶)}

second dividing means for dividing a signal of sub-frames into a first signal (e.g., segment (1)) having a length m, where m is an integer and m<L, where L is the length of said sub-frame, and a second signal (e.g., segment (1')) having a length (L-m) as a reference signal; (p.II-556, Fig.2 & col.1)
{The segmentation of the signal is based on window length N, where N is an integer (Fig.3). The length of segments (1), (1'), (2), (2') are integers.}

finding means for finding a signal (e.g., segment (2)) closest to said reference signal (segment (1')) from an other sub-frame, (p.II-556, col.1)
{WSOLA technique attempts to find a signal, e.g., segment (2), that resembles the reference segment (1') as closely as possible based on the cross-correlation between segment (2) and (1').}

modifying means shortens the signal (e.g., x(n)) by concatenating the first signal (e.g., segment (1)) and the closest signal (segment (2)). (p.II-556, Fig.2 & col.1)

	{The concatenated signal, y(n), is the result of shortening the signal (x(n)) by generating segment (a) from (1) and segment (b) from (2), and by the process of overlap-add segments (a) and (b).} It would have been obvious to a person of ordinary skill in the art at the time the invention was made to modify the excitation source modifying means of Kleijn to include the waveform overlap-add technique of Verhelst in order to shorten the predictive residual signal by concatenating the first signal and closest reference signal. The overlap-add technique provides a concatenated signal that maximizes the similarity to the original waveform and is computationally efficient. Furthermore, these features are important in application such as voice-mail and dictation-tape playback where the speaking rate is a factor that can affect the intelligibility of the playback speech. (see Abstract and Introduction, p.II-554, col.1)
Claim(s) 2	Canceled.
Claim(s) 3	<u>The combination of Kleijn and Verhelst shows:</u> the audio signal processing apparatus as set forth in claim 1, wherein said finding means calculates cross-correlation values with said reference signal (e.g., segment (1)) for signals of said other sub-frame, takes out a signal (e.g., segment (2)) as the closest signal from a position where the calculated cross-correlation value becomes the largest (e.g., maximum). (Verhelst: p.II-556, col.1, last ¶, ll.11-14)
Claim(s) 4	<u>The combination of Kleijn and Verhelst shows:</u> the audio signal processing apparatus as set forth in claim 1, wherein said finding means calculates a square error with said reference signal for signals of said other sub-frame and takes out a signal as the closest signal from a position where the calculated square error becomes the smallest. (Verhelst: see the normalized cross-correlation coefficient equation on p.II-556, col.2, last ¶, ll.11-14) *{The divisor of the normalized cross-correlation coefficient equation is the calculated square error which becomes the smallest when the cross-correlation is maximized.}*
Claim(s) 5	<u>Kleijn shows:</u>

	The audio signal processing apparatus as set forth in claim 1, wherein said excitation source modifying means extends said predictive residual signals by a predetermined extension rate by finding a signal having a predetermined length (e.g., one-pitch period) from the end of the predictive residual signals of a frame; and (col.4, L.61-65; col.8, L.1-24) *{For successive prototype waveform extraction, the current residual signal segment is extended by one pitch period from the end (or center) of the previous residual signal segment to the end (or center) of the current residual signal segment respectively. (col.4, L.61-65; col.8, L.1-24; Fig.12a; col.9, L.37-62)}* concatenating said signal after the end of the predictive residual signals to generate extended predictive residual signals (reconstructed residual). (col.2, L.45-47) *{Each prototype waveform extracted from the residual signal is a representative of a residual signal segment. Concatenation of the successive prototype waveforms generates the reconstructed predictive residual signal. (col.2, L.45-47)}*
Claim(s) 6	Kleijn shows: The audio signal processing apparatus as set forth in claim 1, wherein said synthesizing means comprises a linear prediction (LP) code synthesis filter **303**. (col.2, L.47-48)
Claim(s) 7	Kleijn shows: An audio signal processing apparatus (speech coding apparatus: col.2, L.36-37) for reproducing an audio signal by decoding encoded predictive (LP residual, Fig.10: 203; Fig.2; col.4, L.19-20) residual signals produced by forward prediction on a frame by frame basis, the apparatus comprising: excitation source modifying means (Fig.10: 231) for extending the predictive residual signals by connecting data estimated by extrapolation to signals of a frame while maintaining the pitch, and (col.2, L.41-42; col.4, L.57-65) *{1. Excitation source modifying means comprise performing pitch detection and extracting prototype waveform from the linear predictive (LP) residual signal. (col.2, L.41-42)*

2. In extracting prototype waveforms, the residual signal segment is extended for at least one-half pitch period. (col.4, L.57-65)
3. Forward prediction is e.g., extending the residual signal (periodic for voiced signal) for prediction using past signal samples.}

synthesizing means (Fig.11: 321, 322, 323, 303) for synthesizing an audio signal based on predictive residual signals converted by said excitation source modifying means, (col.2, L.42-L.48)

Kleijn does not show:

wherein said excitation source modifying means comprises:

dividing means for dividing a signal of said sub-frame into a first signal having a length m, where m is an integer and m<L, where L is the length of said sub-frame, and a second signal having a length, L-m, as a reference signal;

finding means for finding a signal closest to said reference signal from an other sub-frame,

wherein said excitation source modifying means shortens said predictive residual signals by concatenating the first signal and the closest signal.

Verhelst teaches:

dividing means for dividing a signal of sub-frames into a first signal (e.g., segment (1)) having a length m, where m is an integer and m<L, where L is the length of said sub-frame, and a second signal (e.g., segment (1')) having a length (L-m) as a reference signal; (p.II-556, Fig.2 & col.1)
{The segmentation of the signal is based on window length N, where N is an integer (Fig.3). The length of segments (1), (1'), (2), (2') are integers.}

finding means for finding a signal (e.g., segment (2)) closest to said reference signal (segment (1')) from an other sub-frame, (p.II-556, col.1)
{WSOLA technique attempts to find a signal, e.g., segment (2), that resembles the reference

	segment (1') as closely as possible based on the cross-correlation between segment (2) and (1').} modifying means shortens the signal (e.g., x(n)) by concatenating the first signal (e.g., segment (1)) and the closest signal (segment (2)). (p.II-556, Fig.2 & col.1) *{The concatenated signal, y(n), is the result of shortening the signal (x(n)) by generating segment (a) from (1) and segment (b) from (2), and by the process of overlap-add segments (a) and (b).}* It would have been obvious to a person of ordinary skill in the art at the time the invention was made to modify the excitation source modifying means of Kleijn to include the waveform overlap-add technique of Verhelst in order to shorten the predictive residual signal by concatenating the first signal and closest reference signal. The overlap-add technique provides a concatenated signal that maximizes the similarity to the original waveform and is computationally efficient. Furthermore, these features are important in application such as voice-mail and dictation-tape playback where the speaking rate is a factor that can affect the intelligibility of the playback speech. (see Abstract and Introduction, p.II-554, col.1)
Claim(s) 8	Canceled.
Claim(s) 9	<u>The combination of Kleijn and Verhelst shows:</u> the audio signal processing apparatus as set forth in claim 7, wherein said excitation source modifying means comprises: first multiplying means for multiplying said reference signal by a first window function; (see Fig.3, p.II-556) *{Segment (1) in Fig.2 is selected by a window function.}* second multiplying means for multiplying signal taken out from said other sub-frame by a second window function; (see Fig.3, p.II-556) and *{Segment (2) in Fig.2 is selected by another window function.}* adding means for adding results of said first and second multiplying means (e.g., adding segment (a) & (b)), (see Fig.2)

	wherein said excitation source modifying means concatenates (e.g., overlaps) results of said adding means after the first signal taken out from said sub-frame to generate one pitch worth of new predictive residual signals. (see Fig.2) *{y(n) is the concantenated result of overlapping and adding segment (a) & (b). y(n) has one pitch period, L_k - L_{k-1}.}*
Claim(s) 10	The combination of Kleijn and Verhelst shows: the audio signal processing apparatus as set forth in claim 7, wherein said finding means calculates cross-correlation values with said reference signal (e.g., segment (1)) for a signal of said other sub-frame and takes out a signal (e.g., segment (2)) as the closest signal from a position where the calculated cross-correlation value becomes the largest (e.g., maximum). (Verhelst: p.II-556, col.1, last ¶, ll.11-14)
Claim(s) 11	The combination of Kleijn and Verhelst shows: the audio signal processing apparatus as set forth in claim 7, wherein said finding means calculates a square error with said reference signal for a signal of said other sub-frame and takes out a signal as the closest signal from a position where the calculated square error becomes the smallest. (Verhelst: see the normalized cross-correlation coefficient equation on p.II-556, col.2, last ¶, ll.11-14) *{The divisor of the normalized cross-correlation coefficient equation is the calculated square error which becomes the smallest when the cross-correlation is maximized.}*
Claim(s) 12	Kleijn shows: The audio signal processing apparatus as set forth in claim 7, wherein said excitation source modifying means extends said predictive residual signals by a predetermined extension rate by finding a signal having a predetermined length (e.g., one-pitch period) from the end of the predictive residual signals of a frame; and (col.4, L.61-65; col.8, L.1-24) *{For successive prototype waveform extraction, the current residual signal segment is extended by one pitch period from the end (or center) of the previous residual signal segment*

	to the end (or center) of the current residual signal segment respectively. (col.4, L.61-65; col.8, L.1-24; Fig.12a; col.9, L.37-62)} concatenating said signal after the end of the prediction residual signals to generate extended predictive residual signals (reconstructed residual). (col.2, L.45-47) *{Each prototype waveform extracted from the residual signal is a representative of a residual signal segment. Concatenation of the successive prototype waveforms generates the reconstructed predictive residual signal. (col.2, L.45-47)}*
Claim(s) 13	Kleijn shows: The audio signal processing apparatus as set forth in claim 7, wherein said synthesizing means comprises a linear prediction (LP) code synthesis filter **303**. (col.2, L.47-48)
Claim(s) 14	Kleijn shows: An audio signal processing method (speech coding method: col.2, L.36-37) for extending or shortening predictive (LP residual, Fig.10: 203; Fig.2; col.4, L.19-20) residual signals on a time axis in decoding a signal encoded by forward prediction on a frame by frame basis, comprising the steps of: *{Speech coding method is for extending predictive residual signals on a time axis.}* processing for extending (Fig.10: 231) the predictive residual signals by connecting data estimated by extrapolation to signals of a frame while maintaining the pitch so as to extend the signals of one frame, and (col.2, L.41-42; col.4, L.57-65) *{1. Excitation source modifying means comprise performing pitch detection and extracting prototype waveform from the linear predictive (LP) residual signal. (col.2, L.41-42)* *2. In extracting prototype waveforms, the residual signal segment is extended for at least one-half pitch period. (col.4, L.57-65)* *3. Forward prediction is e.g., extending the residual signal (periodic for voiced signal) for prediction using past signal samples.}* processing (Fig.11: 321, 322, 323, 303) for synthesizing (through the LP synthesis filter) an audio signal based on said shortened or extended predictive residual signals (the

reconstructed residual signal), (col.2, L.42-L.48)

Kleijn does not show:

wherein the step of shortening said predictive residual signals includes:

dividing a signal of said sub-frame into a first signal having a length m, where m is an integer and m<L, where L is the length of said sub-frame, and a second signal having a length, L-m, as a reference signal;

finding a signal closest to said reference signal from an other sub-frame; and

concatenating the first signal and the closest signal.

Verhelst teaches:

dividing a signal of sub-frames into a first signal (e.g., segment (1)) having a length m, where m is an integer and m<L, where L is the length of said sub-frame, and a second signal (e.g., segment (1')) having a length (L-m) as a reference signal; (p.II-556, Fig.2 & col.1)
{The segmentation of the signal is based on window length N, where N is an integer (Fig.3). The length of segments (1), (1'), (2), (2') are integers.}

finding a signal (e.g., segment (2)) closest to said reference signal (segment (1')) from an other sub-frame, (p.II-556, col.1)
{WSOLA technique attempts to find a signal, e.g., segment (2), that resembles the reference segment (1') as closely as possible based on the cross-correlation between segment (2) and (1').}

concatenating the first signal (e.g., segment (1)) and the closest signal (segment (2)). (p.II-556, Fig.2 & col.1)
{The concatenated signal, y(n), is the result of shortening the signal (x(n)) by generating segment (a) from (1) and segment (b) from (2), and by the process of overlap-add segments (a) and (b).}

It would have been obvious to a person of ordinary skill in the art at the time the

	invention was made to modify the excitation source modifying means of Kleijn to include the waveform overlap-add technique of Verhelst in order to shorten the predictive residual signal by concatenating the first signal and closest reference signal. The overlap-add technique provides a concatenated signal that maximizes the similarity to the original waveform and is computationally efficient. Furthermore, these features are important in application such as voice-mail and dictation-tape playback where the speaking rate is a factor that can affect the intelligibility of the playback speech. (see Abstract and Introduction, p.II-554, col.1)
Claim(s) 15	Canceled.
Claim(s) 16	<u>The combination of Kleijn and Verhelst shows:</u> the audio signal processing method as set forth in claim 14, further comprising shortening said predictive residual signals by first multiplication processing for multiplying said reference signal by a first window function; (see Fig.3, p.II-556) *{Segment (1) in Fig.2 is selected by a window function.}* second multiplication processing for multiplying a signal taken out from said other sub-frame by a second window function; (see Fig.3, p.II-556) and *{Segment (2) in Fig.2 is selected by another window function.}* adding processing for adding results of said first and second multiplying means (e.g., adding segment (a) & (b)), (see Fig.2) concatenating (e.g., overlapping) the results of said adding processing after the first signal taken out from said sub-frame to generate one pitch worth of new predictive residual signals. (see Fig.2) *{y(n) is the concantenated result of overlapping and adding segment (a) & (b). y(n) has one pitch period, L_k - L_{k-1}.}*
Claim(s) 17	<u>Kleijn shows:</u> The audio signal processing method as set forth in claim 14, further comprising:

	extending said predictive residual signals by a predetermined extension rate by finding a signal having a predetermined length (e.g., one-pitch period) from the end of the predictive residual signals of a frame; and (col.4, L.61-65; col.8, L.1-24) *{For successive prototype waveform extraction, the current residual signal segment is extended by one pitch period from the end (or center) of the previous residual signal segment to the end (or center) of the current residual signal segment respectively. (col.4, L.61-65; col.8, L.1-24; Fig.12a; col.9, L.37-62)}* concatenating said signal at the end of the predictive residual signals to generate extended predictive residual signals (reconstructed residual). (col.2, L.45-47) *{Each prototype waveform extracted from the residual signal is a representative of a residual signal segment. Concatenation of the successive prototype waveforms generates the reconstructed predictive residual signal. (col.2, L.45-47)}*

Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Tim Lao whose telephone number is 703-305-8955. The examiner can normally be reached on M-F, 8:30am-5pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Doris To can be reached on 703-305-4827. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Tim Lao
Examiner
Art Unit 2655

TL
07/26/04

W. R. YOUNG
PRIMARY EXAMINER